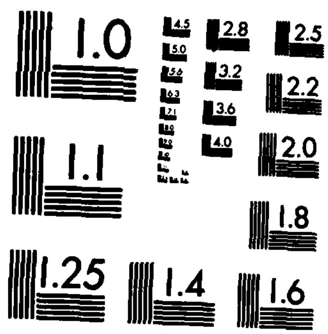
AD-A147 052 TECHNIQUES FOR VECTOR QUANTIZATION(U) CALIFORNIA UNIV SANTA BARBARA DEPT OF ELECTRICAL AND COMPUTER ENGINEERING A GERSHO AUG 84 AFOSR-TR-84-0897 1/1

UNCLASSIFIED AFOSR-82-0008 F/G 12/1 NL

END


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0

AFOSR-TR- 84-0897

4

AD-A147 052

TECHNIQUES FOR VECTOR QUANTIZATION

Final Report for Grant AFOSR 82-0008

October 1, 1982 to September 30, 1983

submitted by

Allen Gersho

Principal Investigator

Department of Electrical and Computer Engineering
University of California
Santa Barbara, California 93106

AFOSR- 82-0008

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 84-0897

6a. NAME OF PERFORMING ORGANIZATION
University of California

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION
Air Force Office of Scientific Research

6c. ADDRESS *(City, State, and ZIP Code)*
Department of Electrical and Computer Engineering, Santa Barbara CA 93106

7b. ADDRESS *(City, State, and ZIP Code)*
Directorate of Mathematical & Information Sciences, AFOSR, Bolling AFB DC 20332

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
AFOSR

8b. OFFICE SYMBOL *(If applicable)*
NM

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR-82-0008

8c. ADDRESS *(City, State, and ZIP Code)*
Bolling AFB DC 20332

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2304	A6	

11. TITLE *(Include Security Classification)*
TECHNIQUES FOR VECTOR QUANTIZATION

12. PERSONAL AUTHOR(S)
Allen Gersho

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(Year, Month, Day)*	15. PAGE COUNT
Final	FROM 1/10/82 TO 30/9/83	AUG 84	10

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

The second year of AFOSR support at the University of California, Santa Barbara, has allowed the investigators to make significant strides in exploring the potential of vector quantization for source coding. Some of this work is described in the attached list of references. Some of the studies reported in the previous report have been completed, including predictive vector quantization and rate distortion modeling of speech using a composite source model to obtain rate distortion bounds on performance of vector quantization. Particularly important results in the second year include the development of a new family of fast search algorithms for pattern matching and the development of Hierarchical Vector Quantization. Several other promising studies, including compandor/lattice coding, were still in progress when the grant terminated.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION

22a. NAME OF RESPONSIBLE INDIVIDUAL
MAJ Brian W. Woodruff

22b. TELEPHONE *(Include Area Code)*
5027

22c. OFFICE SYMBOL

84 10 23 188

1. Summary

The second year of AFOSR support at the University of California, Santa Barbara ~~(UCSB)~~ has allowed us to make significant strides in exploring the potential of vector quantization for source coding. Some of this work is described in the attached list of references. Some of the studies ~~reported in the previous report [12]~~ ~~have been~~ were completed, including predictive vector quantization and rate distortion modeling of speech using a composite source model to obtain rate distortion bounds on performance of vector quantization. Particularly important results in the second year include the development of a new family of fast search algorithms for pattern matching and the development of Hierarchical Vector Quantization. Several other promising studies, including compandor/lattice coding, were still in progress when the grant terminated.

2. Objective

Vector Quantization (VQ) is the constructive implementation of the block source coding idea used by Shannon to obtain his fundamental results on the ultimate rate-distortion performance theoretically achievable. The goal of this project was to explore constructive techniques for VQ that will circumvent the exponential growth of complexity with dimensionality (block size) that is associated with block source coding. The intent is to develop new tools and algorithms that will significantly advance the state of the art in efficient digitization of analog signal sources. The experimental evaluation of new techniques has focused primarily on speech signals although we have also been applying some of our ideas to image coding.

3. Introduction

In the first year of this project, a major start-up effort was needed to establish a viable research program in VQ. A new laboratory and a new computer facility were established to couple theoretical work with experimental computer simulation studies. Thanks to the AFOSR grant, this transition was quite rapid and effective and a very active research program in VQ has been established at UCSB. Several students have been involved in the program, including two senior graduate students as research assistants.

The application of our work to speech coding provides a benchmark to compare the effectiveness of VQ with well established and highly evolved conventional methods for speech coding. In this way, we have been able to realistically assess the impact and effectiveness of VQ techniques.

The next few sections briefly describe the major research accomplishments resulting from the AFOSR support.

4. Fast Search Algorithms for Pattern Matching

Given an optimal codebook for VQ, generated by a clustering algorithm based on (random) training data, it has been generally assumed by workers in vector quantization that the usual exhaustive search encoding is needed to find the nearest neighbor codeword for a given input vector. The resulting complexity depends exponentially on dimensionality so that a dimensionality beyond approximately 10 becomes technologically infeasible to implement. To avoid this severe limitation, suboptimal codebook designs have been proposed that have special structures which avoid the need for exhaustive search but result in a significant performance degradation. In particular, lattice codebooks, multistage quantizers, and suboptimal trees resulting from hierarchical cluster-designs are examples of suboptimal quantizer designs.

We have recognized that it is in fact *not* necessary to compromise optimality in order to reduce the encoding complexity. See [1]. The basic idea is that a given codebook determines an optimal Voronoi partition of the space into nearest neighbor regions. Each such region is defined by an intersection of hyperplanes and by suitable precomputation it is possible to develop an exact and optimal tree structure for searching through *any* given codebook where each branch point in the tree requires a single comparison between two candidate codevectors. All subsequent searching stages then depend on the outcome of previous stages. The design of this optimal tree structure is based on the fact that each hyperplane is defined by two codevectors and that only a small subset of all possible hyperplanes actually determine a particular Voronoi cell. Reference [1] describes an approach that eliminates "remote" hyperplanes.

Most, or in some cases all, of the hyperplanes defining a particular Voronoi cell are determined by the *touching points*, the minimum distance neighboring codevectors, of a particular cell. The number of hyperplanes defining each cell of a code determines the minimum complexity needed to encode a given vector (block) by an optimal vector quantizer (source code). In order to understand the fundamental limits of complexity needed to perform optimal coding, we studied the average number of touching points of an arbitrary code and obtained a sharp upper bound. This result is reported in Ref. [2].

We have been exploring the design of efficient algorithms for generating an optimal encoding tree for a given codebook. The computation to generate such a tree is an "off-line" preprocessing task, which once done can then be used for all real-time encoding. However, it is nevertheless an extremely high complexity task because the optimal tree design requires knowledge of not only the hyperplanes bounding each Voronoi cell but also the vertices of each cell.

We have found a very efficient approach to optimal tree design and have implemented software for this design. Tests on the performance of this algorithm have been entirely successful. Specifically, given several codebooks we compared the encoding performance of the optimal tree and the exhaustive search encoding for each and found no degradation in distortion obtained with the tree structured coder. Of course, the computational complexity of encoding was greatly reduced. In general this approach reduces encoding complexity from N to approximately logN. The algorithm for the precomputation of the optimal tree is an extremely high complexity one and is not pratical for higher dimensional codebooks with the available computing resources. This work has been reported in [7].

The promise of circumventing the exponential complexity of codebook searching if a suitable precomputation can be made, led us to explore alternate approaches to fast search algorithms where a more reasonable precomputation will lead to suitable data structures for a given codebook to support a high speed real time VQ encoder. This study resulted in several new algorithms, all based on a geometric approach to identifying in which Voronoi cell a particular input vector is contained. One method retains the goal of finding a near optimal tree structure, but uses the Voronoi cells associated with the training set to find the tree rather than using the true Voronoi cells associated with the resulting codebook. As a result, the precomputation is a moderately low complexity task and is readily performed for higher dimensional VQ applications. The performance of the resulting fast search algorithm is only very slightly suboptimal in the sense that input space is now partitioned into regions that are not identical to the Voronoi cells associated with the true codebook. Of course, the 'true' codebook is obtained using a cluster-based design from the training set. Hence, the training set is the more fundamental starting point, since it describes the source statistics and it is used to generate the codebook.

In addition to the tree search methods, we have developed three alternate techniques, the Projection Method, the Hyperplane Method, and the Minimax Method for fast pattern matching. Each allows a drastic reduction in encoding complexity for VQ by greatly reducing the number of multiplications while increasing the number of comparison operations. These algorithms have been reported in [8].

We believe this work is a significant and new direction in source coding. It should also have major importance in the field of pattern recognition where the complexity of nearest-neighbor encoding is a fundamental limitation on the capability of template matching recognition systems.

5. Predictive Vector Coding

We have recognized that to counter the complexity problem, VQ must be combined with other redundancy removal techniques. In particular, vector prediction can be used to exploit the correlation between consecutive vectors in a vector source. Once the inter-vector redundancy is reduced, then vector quantization can be used to remove the intra-vector redundancy. Results using this approach were described in last year's AFOSR final report [12].

This work is described in Reference [3] and with greater depth in [4]. A more comprehensive journal paper on the approach has been submitted for publication [9]. The major result of this study was that a performance comparable to the best obtainable by any waveform speech coder at the rate 16 Kb/s has been obtained for the first time using VQ. Thus, we have shown for the first time ever, that VQ has the potential of being a competitive technique for waveform coding of speech. We have presented a tutorial-survey of the use of VQ for waveform coding in [11, 12].

6. Rate Distortion Bounds for Speech Coding

In the first study of waveform coding of speech using vector quantization, Abut and Gray [6] attempted to compare the actual attained rate-distortion with that predicted from a Markov model of the speech signal. Their experimental results showed that the actual performance obtained considerably *exceeded* the theoretically best performance predicted by rate distortion theory. This apparent contradiction to the theoretical limitation predicted by rate distortion theory is explained by the fact that the speech model used was grossly inadequate. We have been exploring the application of our new techniques for VQ of speech waveforms and have felt the need for a more meaningful model of speech so that useful comparisons with rate distortion theory can be made. We have developed a new model for describing speech and for experimentally generating a speech-like waveform for simulation purposes. This work was reported in last year's final report [12]. A paper on this work is in preparation.

7. Hierarchical Vector Quantization

We have recently introduced a technique for coding extremely high dimensionality vectors that could represent linguistically distinct segments of an information source. By partitioning such a "supervector" into a large number of minivectors and extracting a scalar feature from each minivector, i.e., the rms energy, we form a lower dimensionality feature vector. This feature vector may still be of too high a dimension for direct vector quantization/classification with a single codebook. Hence, we repeat the process and partition this feature vector into subvectors, extract a feature from each subvector, and form a higher level (but lower dimensionality) vector of features. This process could continue to yet higher levels, but it is often sufficient at this third stage to allow direct vector quantization of the upper level feature vector. The subvectors at each level are then each coded using dynamic codebook allocation determined by the next higher level. At each level, a fixed total number of bits is optimally allocated to efficiently code all subvectors at that level with the varying precision that the individual subvectors require. Finally, each of the minivectors of the original waveform is then coded with another set of codebooks for this level, again with dynamic codebook allocation.

This hierarchical approach has promising implications for efficiently coding linguistically distinct source units for recognition, classification, or compression applications. Our best quality speech compression so far has been obtained with this approach. We have developed a novel speech segmentation algorithm for this application. Further work is continuing since the termination of the AFOSR grant to find ways to jointly optimize the family of codebooks at different levels, to find optimal codebook allocation algorithms, and to incorporate perceptual weighting strategies into the design and coding procedures. Two papers describing aspects of this work that were supported in part by the AFOSR grant are [13] and [14].

8. Concluding Remarks

Vector quantization has been found to be more than a theoretical curiosity and is now recognized as a major and viable technique for both speech and image compression. As digital processing and memory technology has advanced, the interest in VQ has become widespread and is now being pursued by numerous researchers in the U.S. and abroad. The AFOSR support has allowed us to play an important pioneering role in exploring the potential of VQ. Although some of our research projects could not be completed due to the termination of AFOSR support in September 1982, several important new techniques were established in our group and are being pursued by other researchers elsewhere. We have been fortunate in obtaining other sources of support for our work and some of the studies initiated under AFOSR support are being continued at this time. We are pleased to have had the opportunity to contribute to this important new direction for data compression.

9. References

[1] A. Gersho, "On the structure of vector quantizers," *IEEE Trans. on Information Theory*, March 1982.

[2] C. Wang and A. Gersho, "A bound on the number of nearest neighbors of a Euclidean code," submitted for publication to *IEEE Trans. on Information Theory*,

[3] V. Cuperman and A. Gersho, "Adaptive vector differential coding of speech," *Conf. Record, IEEE Global Telecommunications Conference*, Miami, Dec. 1982.

[4] V. Cuperman, "Efficient Waveform Coding Using Vector Quantization," Ph.D. dissertation UCSB March 1983, in preparation.

[5] A. Gersho and B. Ramamurthi, "Image Coding Using Vector Quantization," *Proc. 1982 Conf. Acoustics, Speech, and Signal Processing*, Paris, May 1982.

[6] H. Abut, R. M. Gray, G. Rebolledo, "Vector Quantization of Speech and Speech-Like Waveforms," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-30, pp. 423-436, June 1982.

[7] A. Gersho and D. Y. Cheng, "Fast Nearest Neighbor Search for Nonstructured Euclidean Codes," *IEEE Intrl. Symp. on Inform. Theory*, September 1983.

* [8] D. Y. Cheng, A. Gersho, B. Ramamurthi, and Y. Shoham, "Fast Search Algorithms for Vector Quantization and Pattern Matching," *Proc. Int'l Conf. Acoustics, Speech, and Signal Processing*, San Diego, March 1984.

[9] V. Cuperman and A. Gersho, "Adaptive Vector Predictive Coding of Speech," *IEEE Trans. on Communications, submitted for publication*.

* [10] Allen Gersho and Vladimir Cuperman, "Speech Waveform Coding Using Vector Quantization," Conf. Record, IEEE Global Communications Conference, December 1983.

* [11] A. Gersho and V. Cuperman, "Vector Quantization: A Pattern Matching Technique for Speech Coding," IEEE Communications Magazine, vol. 21, no. 9, pp. 15-21, December 1983.

[12] A. Gersho, P.I., "Techniques for Vector Quantization," Final Report, AFOSR Grant 82-0008, October 1, 1981 to September 30, 1982.

* [13] Allen Gersho and Yair Shoham, "Hierarchical Vector Quantization of Speech with Dynamic Codebook Allocation," *Proc. Int'l Conf. Acoustics, Speech, and Signal Processing*, San Diego,

March 1984.

* [14] Y. Shoham and A. Gersho, "Pitch Synchronous Transform Coding of Speech at 9.6 Kb/s Based on Vector Quantization," Conf. Rec., *1984 IEEE Inter. Conf. on Communications*, Amsterdam, May 1984, pp. 1179-1182.

[15] Allen Gersho and Vladimir Cuperman, "Vector Predictive Source Coding" *IEEE Intrl. Symp. on Inform. Theory*, September 1983.

* Copy appended to this report.

END

FILMED

11-84

DTIC